

On the photodissociation of H_2 by the first stars

S.C.O. Glover & P.W.J.L. Brand

Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ

1 June 2000

ABSTRACT

The first star formation in the universe is expected to take place within small protogalaxies, in which the gas is cooled by molecular hydrogen. However, if massive stars form within these protogalaxies, they may suppress further star formation by photodissociating the H_2 . We examine the importance of this effect by estimating the timescale on which significant H_2 is destroyed. We show that photodissociation is significant in the least massive protogalaxies, but becomes less so as the protogalactic mass increases. We also examine the effects of photodissociation on dense clumps of gas within the protogalaxy. We find that while collapse will be inhibited in low density clumps, denser ones may survive to form stars.

Key words: galaxies:formation

1 INTRODUCTION

Although we can now detect star-forming galaxies at very high redshifts, the ionization and ubiquitous metal enrichment of the IGM at these redshifts suggest that we are not yet seeing the very earliest star formation. In order to study the formation of the first stars, we are thus forced to rely on a purely theoretical approach. Work over the past few decades has led to substantial progress in our understanding of this epoch, and while many of the details are unclear, a broad outline of events can be constructed.

The first step is the gravitational collapse of overdense regions in the early universe. In cold dark matter models, this happens hierarchically, with objects on the smallest scales forming first. In order for gas to collapse along with the dark matter, however, gravitational forces must be strong enough to overcome the thermal pressure of the gas. This occurs for overdensities more massive than the Jeans mass, M_J , which, prior to reionization, is of the order of $10^4 M_\odot$ (Haiman, Abel & Rees 1999).

As the gas collapses, it is heated by adiabatic compression and by shocks. In the absence of cooling, this leads to an increase in pressure which will eventually halt the collapse. Moreover, it prevents fragmentation of the gas. In order for star formation to occur, continued collapse and fragmentation are required, and thus stars will form only if the gas can lose heat effectively. In quantitative terms, it is necessary that the cooling timescale be shorter than the free-fall timescale (Rees & Ostriker 1977).

During the collapse of the first protogalaxies, the gas temperature typically reaches only a few thousand Kelvin (Tegmark et al. 1997). This is significantly lower than the temperature required for the gas to be able to cool via Lyman- α emission. In the absence of metals, the only effective low temperature coolant is molecular hydrogen. This

forms mainly via the reactions



although a small amount also forms via



While the final H_2 fraction is generally small, it is nonetheless able to cool primordial gas effectively, as has been established by a number of authors (see Haiman, Thoul & Loeb 1996 and references therein).

The subsequent evolution of the gas and the eventual formation of stars are still not well understood, although substantial progress in this area has recently been made (Abel, Bryan & Norman 2000; Bromm, Coppi & Larson 1999; Nakamura & Umemura 1999).

Finally, once star formation begins, it will influence the surrounding gas via a number of feedback processes. Two processes in particular have been studied: energy input from supernovae and UV radiation from stars.

Since the energy of a single supernova is comparable to the binding energy of the first protogalaxies, it is highly probable that they will be entirely disrupted by supernovae (MacLow & Ferrara 1999; Ferrara & Tolstoy 1999). However, as this occurs only once the first massive stars reach the end of their life, it is ineffective during the first few million years of star formation. Moreover, the effects of supernovae are quite local, being restricted to a single protogalaxy and its immediate surroundings.

Ultraviolet radiation, on the other hand, can potentially act much faster, and over a much greater volume. It exerts a feedback by photodissociating H_2 , thereby suppressing cooling. The global effects of ultraviolet radiation from the first stars have recently been studied by Haiman et al. (1999) and

Ciardi et al. (1999). They show that a soft ultraviolet background builds up prior to reionization, which photodissociates H_2 in the IGM and in newly collapsing protogalaxies.

Less work has been done on the local effects of UV radiation, i.e. the ability of the first massive stars to suppress H_2 cooling within their host protogalaxy. This has recently been studied by Omukai & Nishi (1999), who find that this local feedback is very effective, with the radiation from a single massive star sufficient to dissociate H_2 throughout a protogalaxy. However, their treatment of the effects of the self-shielding of H_2 is correct only once the photodissociation region created by the star has reached equilibrium. Without knowing the timescale on which this occurs, it is impossible to assess the true effectiveness of the local feedback from UV radiation.

In this paper we overcome this problem by introducing a simple method to estimate the photodissociation timescale. We first calculate the rate at which photodissociating photons are produced by a primordial star, approximated as a blackbody. Next, we calculate how many of these photons illuminate the object of interest, how many are absorbed, and finally, how many of these absorptions lead to photodissociation. Most of these quantities can be estimated to within a factor of a few or better; even in the worst case, we expect our final estimate to be inaccurate by an order of magnitude at most.

Applying this method to simple protogalactic models, we find that the photodissociation timescale depends strongly upon the mass of the protogalaxy. In the least massive protogalaxies, equilibrium is reached on a timescale that is short compared to the stellar lifetime, while in larger protogalaxies, the two timescales are comparable. We also examine the fate of H_2 in dense clumps of gas within the protogalaxy. Such clumps are the likely precursors of star formation, but can collapse only if enough H_2 survives within them to provide sufficient cooling. We find that this is only the case within the densest clumps.

We thus partially confirm the results of Omukai & Nishi. Ultraviolet feedback is very effective at suppressing cooling in small protogalaxies, if the amount of gas in dense clumps is small. It is less effective in larger systems, or when much of the gas is densely clumped. This suggests that UV feedback alone need not imply a low star formation efficiency.

In section 2 we examine the manner in which UV radiation leads to the photodissociation of H_2 , and briefly discuss the approach taken by Omukai & Nishi and the importance of a proper treatment of self-shielding. In section 3 we introduce our approximate model, and in section 4 apply it to our protogalactic model. In section 5 we examine the effects of a number of complications not included in our simple model, and we conclude in section 6.

2 PHOTODISSOCIATION OF H_2

In the cool primordial gas that forms the first protogalaxies, the vast majority of the H_2 will lie in its lowest possible electronic and vibrational energy state, with much also lying in the rotational ground state. In order to destroy ground-state H_2 molecules, a substantial energy input is required. At high temperatures this can be provided by collisions, but at the temperatures typically found within the first protogalaxies,

collisional dissociation is unimportant. In this case H_2 is destroyed mainly by radiative reactions. There are several possibilities.

- (i) Photoionization:



(O’Neil & Reinhardt 1978). This has a threshold of 15.42 eV, well above the energy required for H I ionization. It will only occur within the H II region surrounding the radiation source. As we show in section 5.3, this region may be quite small.

- (ii) Excitation to the vibrational continuum of an excited electronic state of H_2 :



(Allison & Dalgarno 1969). This has a threshold of 14.16 eV for ortho-hydrogen and 14.68 eV for para-hydrogen. This too is restricted to H II regions.

- (iii) Direct excitation to the vibrational continuum of the ground electronic state of H_2 . This process is strongly forbidden and proceeds at a negligible rate.

- (iv) Two-step photodissociation (the Solomon process):



(Stecher & Williams 1967). The first step involves excitation to either the $B^1\Sigma_u^+$ or $C^1\Pi_u$ excited electronic states, with threshold energies of 11.15 eV and 12.26 eV respectively. This is followed by decay to some vibrational level of the ground electronic state. A fraction of these decays will be to the vibrational continuum of the ground state, resulting in dissociation. A number of rotational and vibrational states of $B^1\Sigma_u^+$ and $C^1\Pi_u$ are available from the ground state by absorption below 13.6 eV; the associated spectral lines are called the Lyman and Werner bands.

The fate of H_2 in cold, neutral gas thus depends upon the flux of UV in the Lyman and Werner bands. Unfortunately, this is difficult to calculate accurately. Direct solution of the radiative transfer problem, even if decoupled from the hydrodynamics, is computationally expensive. It is thus useful to explore analytic or semi-analytic alternatives which can give reasonable estimates of the effects of the radiation.

One possible technique is that suggested by Omukai & Nishi. They consider two processes: formation of H^- by radiative association (reaction 1 above), which is the rate determining step in the formation of H_2 , and two step photodissociation. The former has a rate coefficient (de Jong 1972)

$$k_{\text{form}} = 1.0 \times 10^{-18} T \text{ cm}^3 \text{ s}^{-1} \quad (8)$$

for gas at a temperature T , while the rate coefficient for the latter can be written, in the optically thin limit, as (Draine & Bertoldi 1996)

$$k_{2\text{step}} = 1.13 \times 10^8 F_{\text{LW}} \text{ s}^{-1}, \quad (9)$$

where F_{LW} is the average flux density in the Lyman-Werner bands (in units of $\text{erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}$). With this very simplified chemistry, H_2 reaches chemical equilibrium on a timescale $t_{\text{dis}} = k_{2\text{step}}^{-1}$. If this is significantly shorter than the timescales on which the flux, temperature, density or

ionization vary then we can solve for the equilibrium fractional abundance of H₂

$$\begin{aligned} x_{\text{H}_2} &= \frac{k_{\text{form}}}{k_{2\text{step}}} x_e n \\ &= 8.8 \times 10^{-27} x_e F_{\text{LW}}^{-1} T n, \end{aligned} \quad (10)$$

where x_e is the fractional ionization and n is the total number density of the gas, in units of cm⁻³.

If self-shielding of the Lyman-Werner bands is unimportant then one can define the ‘region of influence’ of a massive star to be the volume of gas surrounding it which satisfies two conditions; first, that the cooling time, given the equilibrium H₂ abundance, is less than the free-fall timescale of the gas, and second, that this equilibrium abundance can be reached in less than the lifetime of the central star. Omukai & Nishi consider the particular example of gas illuminated by a massive star with average luminosity density in the Lyman-Werner bands of $L_{\text{LW}} = 10^{24}$ erg s⁻¹ Hz⁻¹. In this case, using their values of temperature, density and ionization, and assuming a uniform distribution of gas, the cooling condition implies a region of influence of radius

$$r_{\text{cool}} = 110 \text{ kpc}, \quad (11)$$

while the equilibrium condition implies a radius of

$$r_{\text{eq}} = 9.4 \text{ kpc}. \quad (12)$$

Both of these values are significantly larger than the typical virial radius of a protogalactic halo,

$$r_{\text{vir}} = 200 \left(\frac{M}{10^6 M_{\odot}} \right)^{1/3} \text{ pc}, \quad (13)$$

(where we have followed Omukai & Nishi in adopting a gas number density $n = 1 \text{ cm}^{-3}$). Hence, Omukai & Nishi conclude that a single massive star can effectively suppress star formation throughout the protogalaxy.

Self-shielding becomes important for gas at a radius r from the star when the intervening H₂ column density, N_{H_2} , exceeds 10^{14} cm^{-2} . In this case the situation is more complicated. The effect of self-shielding is to attenuate the Lyman-Werner flux. For the averaged flux this can be modelled by a shielding factor f_{sh} , given with reasonable accuracy by (Draine & Bertoldi 1996)

$$f_{\text{sh}} = \min \left[1, \left(\frac{N_{\text{H}_2}}{10^{14}} \right)^{-0.75} \right]. \quad (14)$$

However, unless the H₂ abundance has already reached equilibrium, then the column density of H₂ between the star and a given point, and hence the Lyman-Werner flux at that point, will change over time. Omukai & Nishi assume that the H₂ abundance has reached equilibrium and use the previously derived equilibrium abundance to calculate the distance from the star at which $N_{\text{H}_2} = 10^{14} \text{ cm}^{-2}$. Beyond this radius, F_{LW} drops sharply, and they consider this the effective boundary of the region of influence of the star. For the same example as above, they find that

$$r_{\text{sh}} = 0.97 \text{ kpc}. \quad (15)$$

This is still comparable to the size of a typical protogalaxy. However, Omukai & Nishi do not calculate the time taken to reach the equilibrium state. This may be significantly longer than in the optically thin case, as in the initially self-shielded

gas F_{LW} and hence $k_{2\text{step}}$ have values much lower than those in the equilibrium case. It is not clear that this equilibrium can be reached within the lifetime of the star.

We attempt to remedy this difficulty by proposing a method for estimating t_{dis} for optically thick gas, based upon a simple calculation of the time taken for the star to emit sufficient photons to dissociate the bulk of the H₂. We outline our method in the next section.

3 ESTIMATING THE DISSOCIATION TIMESCALE

For a cloud containing a mass M_{H_2} of H₂, we define the photodissociation timescale as

$$t_{\text{dis}} = \frac{M_{\text{H}_2}}{|\dot{M}_{\text{H}_2}|}, \quad (16)$$

where \dot{M}_{H_2} is the mass of H₂ destroyed per second. In terms of the number of H₂ molecules destroyed per second, \dot{N}_{H_2} this is simply

$$\dot{M}_{\text{H}_2} = -m_{\text{H}_2} \dot{N}_{\text{H}_2}, \quad (17)$$

where m_{H_2} is the mass of a hydrogen molecule. Now, \dot{N}_{H_2} can be written as the product of four numbers

$$\dot{N}_{\text{H}_2} = N_{\text{dis}} f_{\text{inc}} f_{\text{abs}} f_{\text{dis}}, \quad (18)$$

where N_{dis} is the number of photons capable of photodissociation produced per second, f_{inc} the fraction of these that are incident on the cloud, f_{abs} the fraction of incident photons that are absorbed in the cloud, and f_{dis} the fraction of absorptions that lead to dissociation. To calculate t_{dis} we thus need only calculate these four numbers. To do this precisely is difficult, but fortunately we can estimate them relatively easily.

3.1 N_{dis}

Any photon with sufficient energy can, in principle, cause the dissociation of an H₂ molecule, although the probability of absorption for those photons with frequencies far from the centre of any Lyman-Werner line will be small. To calculate N_{dis} , therefore, we simply calculate the total number of photons between 11.15 eV (the threshold for excitation to $B^1\Sigma_u^+$ from the ground state) and the Lyman limit at 13.6 eV. Higher energy photons will be blocked by neutral hydrogen, and are important only within the H II region.

For a single star, we have

$$N_{\text{dis}} = \int_{\nu_1}^{\nu_2} \frac{L_{\nu}}{h\nu} d\nu, \quad (19)$$

where L_{ν} is the luminosity density of the star, $h\nu_1 = 11.15 \text{ eV}$ and $h\nu_2 = 13.6 \text{ eV}$. For primordial stars, the metallicity is effectively zero and we can treat the stars as black bodies to a good approximation (Cojazzi et al. 1999). In this case

$$L_{\nu} = 4\pi R_{\star}^2 F_{\nu} \quad (20)$$

for a star of radius R_{\star} , where $F_{\nu} = \pi B_{\nu}$ is the flux at frequency ν , B_{ν} is the Planck function

Figure 1. The number of photons with energies between 11.15 eV and 13.6 eV emitted per second by a metal-free star, as a function of stellar mass. The luminosities and effective temperatures of the stars are taken from Cojazzi et al. (1999).

$$B_\nu = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT_{\text{eff}}} - 1}, \quad (21)$$

and the stellar effective temperature is T_{eff} . We can write this in terms of the total luminosity L_* as

$$L_\nu = \frac{L_*}{\sigma T_{\text{eff}}^4} \pi B_\nu. \quad (22)$$

N_{dis} is then given by

$$N_{\text{dis}} = \frac{L_*}{\sigma T_{\text{eff}}^4} \int_{\nu_1}^{\nu_2} \frac{\pi B_\nu}{h\nu} d\nu. \quad (23)$$

We obtain values of L_* and T_{eff} as a function of stellar mass from Cojazzi et al. (1999), who fit an isochrone to a collection of ZAMS metal-free stellar models. It is then simple to calculate N_{dis} as a function of stellar mass. We plot our results in figure 1.

We do not incorporate the effects of stellar evolution into our model, as we expect them to be a relatively small correction.

3.2 f_{inc}

The fraction of the emitted flux incident on a region of interest depends upon the size of the region and its position with respect to the star. For example, for a star in the centre of a spherically symmetric cloud of gas, clearly $f_{\text{inc}} = 1$. On the other hand, for a star illuminating a compact cloud from a distance, $f_{\text{inc}} = \Omega/4\pi$, where Ω is the solid angle subtended by the cloud, as seen from the star.

3.3 f_{abs}

The fraction of incident photons absorbed in the cloud is obviously frequency dependent; far more photons are absorbed at frequencies corresponding to the centres of the Lyman-Werner lines than in the wings. However, rather than tackle this frequency dependence directly, we instead choose to approximate the effects of absorption in the set of Lyman-Werner lines by

$$f_{\text{abs}} \simeq \frac{W_{\text{tot}}}{W_{\text{max}}}, \quad (24)$$

where W_{tot} is the total dimensionless equivalent width of the set of lines and $W_{\text{max}} = \ln\left(\frac{1110}{912}\right) \simeq 0.2$ is the dimensionless

width of the range of wavelengths that contains all of the Lyman-Werner lines.

The dimensionless equivalent width of a single line i is

$$W_i = \int \frac{d\nu}{\nu} (1 - e^{-\tau_\nu}), \quad (25)$$

where τ_ν is the optical depth at frequency ν , and the integral is performed over the line profile. To calculate W_i , we assume that the line shape is well fit by a Voigt profile. We then need only a few parameters to calculate the equivalent width. We take the oscillator strength of the transition and the frequency at line centre from Abgrall & Roueff (1989) and the intrinsic broadening of the upper lines from Abgrall et al. (1992). We also need to know the Doppler broadening parameter,

$$b = \sqrt{\frac{2kT}{m_{\text{H}_2}} + v_t^2}, \quad (26)$$

where v_t is the microturbulent velocity. Clearly, we will not in general know the appropriate value of b to use; however, it is possible to place some limits on it. Our lower limit, $b_{\text{min}} = 1.3 \text{ km s}^{-1}$, corresponds to pure thermal broadening at a temperature of 200 K. This is approximately the lowest temperature at which H_2 cooling is effective; the amount of gas cooler than this is expected to be negligible, unless HD cooling should prove to be important. We obtain an approximate upper limit by assuming thermal broadening at 6000 K (a higher temperature implies significant collisional dissociation), plus a similarly-sized contribution from microturbulence. This gives $b_{\text{max}} \simeq 10 \text{ km s}^{-1}$, and is a strong upper limit. Given some value of b within this range, it is then straightforward to calculate W_i as a function of N_{H_2} by means of the accurate numerical approximations given in Rodgers & Williams (1974).

For a set of lines, one might naively expect that

$$W_{\text{tot}} = \sum_i W_i, \quad (27)$$

ie. that the total dimensionless equivalent width is simply the sum of the widths of the individual lines. Unfortunately, this expression behaves unphysically in the high column density limit, as it does not account for the effects of line overlap. Clearly $W_{\text{tot}} \leq W_{\text{max}}$ – the cloud cannot absorb more light than is actually there. To correct for the effects of overlap we follow the procedure suggested in Draine & Bertoldi (1996), using a modified equivalent width defined as^{*}

$$\tilde{W}_{\text{tot}} = W_{\text{max}} [1 - \exp(-W_{\text{tot}}/W_{\text{max}})]. \quad (28)$$

It is clear from this that $\tilde{W}_{\text{tot}} \rightarrow W_{\text{max}}$ as $W_{\text{tot}} \rightarrow \infty$. We then have

$$f_{\text{abs}} = \frac{\tilde{W}_{\text{tot}}}{W_{\text{max}}} = 1 - \exp(-W_{\text{tot}}/W_{\text{max}}). \quad (29)$$

Finally, to calculate \tilde{W}_{tot} correctly, we must know which lines to include. As previously noted, we expect almost all of the H_2 to be in its ground state. The ortho- to para-hydrogen

^{*} Note that while Draine & Bertoldi also included the effects of absorption in the hydrogen Lyman series in their calculation of the effects of overlap, it is clear from their figure 6 that these generally have a very small effect on \tilde{W}_{tot} .

Figure 2. The fraction of radiation in the Lyman-Werner bands that is absorbed by the gas, for Doppler parameter $b = b_{\min}$, for an ortho- to para- ratio of zero (solid line) or three (dashed line).

Figure 3. The same as figure 2, but with $b = b_{\max}$

ratio is not known *a priori*, but there is good reason to believe that para-hydrogen will predominate (Abel et al. 1997). Assuming that there is *no* ortho-hydrogen, it is easy to calculate \tilde{W}_{tot} , and hence f_{abs} , as a function of N_{H_2} . We plot this in figure 2 for $b = b_{\min}$ and in figure 3 for $b = b_{\max}$. For comparison, in both of these figures we also plot f_{abs} obtained with the assumption of an ortho- to para- ratio of three to one.

We see that uncertainties in the ortho- to para- ratio will affect our calculated value of f_{abs} for all but the largest or smallest column densities. However, the error introduced by this is relatively small – the difference between our two extreme cases is never more than a factor of two. The uncertainty in b is potentially more important, but its effects are restricted to the square-root portion of the curve of growth, between $N_{\text{H}_2} = 10^{14} \text{ cm}^{-2}$ and $N_{\text{H}_2} = 10^{18} \text{ cm}^{-2}$. Even in this region, the difference is only a factor of five and outside of this region it quickly becomes unimportant.

3.4 f_{dis}

The fraction of excitations that are followed by dissociation depends upon the vibrational state of the excited level; values for the individual levels are given in Abgrall et al. (1992). Averaged over all UV absorptions, Draine & Bertoldi find a typical value $f_{\text{dis}} = 0.15$, with a slight dependence on N_{H_2} .

There is, however, a complication. Those absorptions not resulting in dissociation result instead in decay to a bound vibrational state. If this state is sufficiently low-lying, the emitted photon will still be energetic enough to cause photodissociation. In low or moderate column density gas, this has a negligible effect – since effectively all of the H₂ lies

in its ground vibrational state, only decays to this state produce photons with a high probability of re-absorption; decays to excited states will result in photons with frequencies at which the gas is still optically thin. Since decay to an excited state is far more likely than decay to the ground state, it follows that the majority of re-emitted photons simply escape from the gas. On the other hand, in very high column density gas this is no longer the case. At high column densities the wings of the absorption lines become optically thick, and eventually the lines themselves overlap. In this case, it is no longer true that the majority of re-emitted photons escape; rather, the majority will eventually be absorbed.

To incorporate this effect, we choose to define f_{dis} somewhat differently from the analogous quantity in Draine & Bertoldi (1996). Rather than considering the probability of dissociation *per absorption*, we consider the probability of dissociation *per photon* given that the photons undergo at least one absorption. Clearly, in the low density limit, photons undergo only one absorption, and the definitions coincide; on the other hand, in the high density limit, f_{dis} approaches unity.

To properly calculate f_{dis} requires a more accurate solution of the radiative transfer problem than is appropriate for our needs. Instead we adopt the simple *ansatz* that

$$f_{\text{dis}} \simeq 0.15 + 0.85f_{\text{abs}}. \quad (30)$$

This has the correct behaviour in the high and low density limits, and the error introduced elsewhere should be reasonably small.

3.5 Calculating the dissociation timescale

Putting together our various estimates, we have:

$$\begin{aligned} t_{\text{dis}} &= \frac{M_{\text{H}_2}}{|\dot{M}_{\text{H}_2}|} \\ &= \frac{x_{\text{H}_2} M}{m_{\text{H}_2} \dot{N}_{\text{H}_2}}, \end{aligned} \quad (31)$$

where

$$\dot{N}_{\text{H}_2} = N_{\text{dis}} f_{\text{inc}} f_{\text{abs}} f_{\text{dis}} \quad (32)$$

Here x_{H_2} is the fractional abundance of H₂ and M is the total mass of the cloud. Combining the two equations, and writing M in units of M_{\odot} , we have

$$t_{\text{dis}} \simeq 2 \times 10^{49} \frac{x_{\text{H}_2} M}{N_{\text{dis}} f_{\text{inc}} f_{\text{abs}} f_{\text{dis}}} \text{ yr}. \quad (33)$$

For the example of a $25M_{\odot}$ star, $N_{\text{dis}} \simeq 10^{48}$ and

$$t_{\text{dis}} \simeq 20 \frac{x_{\text{H}_2} M}{f_{\text{inc}} f_{\text{abs}} f_{\text{dis}}} \left(\frac{N_{\text{dis}}}{10^{48}} \right)^{-1} \text{ yr}. \quad (34)$$

4 APPLICATIONS

As an example, we apply our method to a very simple protogalactic model, consisting of two components: a spherically symmetric, diffuse component (the halo), in which are embedded clumps of a significantly denser component. Although recent numerical simulations offer some support for this picture, doubts inevitably remain concerning the general applicability of the results. Moreover, observations of

the local molecular clouds suggest that it may be an oversimplification. Indeed, there is evidence that molecular clouds have a fractal structure (Elmegreen 1997), although this remains controversial (Blitz & Williams 1997). Nevertheless, our model has the advantage of simplicity, and should be a reasonable guide.

4.1 The spherical halo

We first consider the simple case of a single star at the centre of a spherical halo. As a representative example, we consider a $25M_{\odot}$ star, which gives $N_{\text{dis}} = 10^{48} \text{ s}^{-1}$, although it is straightforward to rescale our results for a different stellar mass. Now, spherical symmetry implies $f_{\text{inc}} = 1$, so it remains only to specify the mass, fractional H_2 abundance and H_2 column density of the halo. From these we can then calculate f_{abs} and f_{dis} .

We can place a number of constraints upon these parameters. The fractional abundance is the most straightforward. A lower limit upon its value comes from the requirement that there be enough H_2 to cool the gas effectively. The required fractional abundance has been estimated by Tegmark et al. (1997) to be approximately 5×10^{-4} . We err on the side of caution and adopt a lower limit of 10^{-4} . An upper limit can be obtained via a method suggested by Nishi & Susa (1999). They note that the maximum abundance of H_2 that can form in a halo is approximately the amount that forms within some critical timescale. This critical timescale is temperature dependent, and is the shortest of the recombination, cooling and dissociation timescales. For low ionization gas, this typically gives a maximum abundance of order 10^{-3} .[†]

We can also constrain the halo mass by a cooling argument, given the redshift at which it forms. This depends upon the cosmological model. A frequently adopted value is $z = 30$, appropriate to the case of a rare halo in an SCDM cosmology. However, values in the range $10 < z < 100$ are plausible (depending upon the rarity of the halo at formation) if we do not limit ourselves to a particular cosmological model.

Given the redshift, a lower limit upon the halo mass M comes from the requirement that the virial temperature of the halo be high enough to form sufficient H_2 for effective cooling. This has also been studied by Tegmark et al. They find that the critical mass is redshift dependent, increasing at low redshifts. In the range that we consider, their minimum value is $M_{\text{min}} = 10^5 M_{\odot}$. Adopting a gas to dark matter ratio $f_b = 0.1$, this corresponds to a minimum gas mass of $10^4 M_{\odot}$.

An upper limit on the mass comes from the fact that halos with virial temperatures greater than 10^4 K will cool effectively via Lyman- α emission. Although it is unclear what role H_2 cooling plays in such halos, and whether it is necessary for star formation, it is often assumed that Lyman- α cooling suffices. As a detailed examination of this issue lies

beyond the scope of this paper, we also make this assumption, and do not consider larger halos. The mass corresponding to a virial temperature of 10^4 K is given by

$$M_{\text{max}} = 2.2 \times 10^9 h^{-1} (1+z)^{-3/2} M_{\odot}, \quad (35)$$

where $100h \text{ km s}^{-1} \text{ Mpc}^{-1}$ is the Hubble constant. In the range of redshifts we consider, this corresponds to a maximum mass of approximately $10^8 M_{\odot}$ and hence a maximum gas mass of $10^7 M_{\odot}$.

Given the mass and H_2 abundance, we next specify the H_2 column density, N_{H_2} , by choosing a halo density profile. The simplest choice is a uniform sphere. In this case the spherical collapse model suggests that the resulting virialized halo will have an overdensity of $\delta = 18\pi^2$ with respect to the cosmological background at the time of collapse. This is equivalent to a total number density of

$$n_{\text{vir}} = 2.1 \times 10^{-5} \left(\frac{\Omega_b h^2}{0.0125} \right) (1+z)^3 \text{ cm}^{-3}, \quad (36)$$

for a neutral gas of primordial composition, and hence an H_2 number density

$$n_{\text{H}_2} = x_{\text{H}_2} n_{\text{vir}}. \quad (37)$$

Since we also know the mass, it is now trivial to solve for the virial radius of the halo, r_{vir} , and hence for N_{H_2} . We find that, for a star at the centre of the halo, the H_2 column density along a radial ray is

$$N_{\text{H}_2} = 4.7 \times 10^{15} x_{\text{H}_2} M^{1/3} \left(\frac{\Omega_b h^2}{0.0125} \right)^{2/3} (1+z)^2 \text{ cm}^{-2}. \quad (38)$$

Unfortunately, the uniform sphere is an unrealistic model; both theory and computer simulation suggest that real protogalactic clouds will be centrally concentrated. To address this, we consider a second model, that of a truncated isothermal sphere (Shapiro, Iliev & Raga 1999). This has a core overdensity $\delta = 1.796 \times 10^4$, a core radius $r_0 = 0.038 r_{\text{vir}}$ and a truncation radius $r_t = 1.108 r_{\text{vir}}$ (corresponding to $r_t = 29.4 r_0$). Over much of its radius, the profile is approximately isothermal. Comparing the two column densities, we find that

$$N_{\text{T}} = 11.3 N_{\text{U}}, \quad (39)$$

where N_{T} and N_{U} are the column densities for the truncated isothermal sphere and uniform sphere respectively. The uniform sphere likely underestimates the column density by an order of magnitude. However, in the truncated isothermal sphere model, a significant fraction of this column density comes from the small, dense core. Outside of the core, the uniform sphere model may give a better estimate. We therefore continue to consider both models, with the expectation that the uniform sphere model will provide us with a strong lower limit on t_{dis} , while the truncated isothermal model should provide a reasonable upper limit.

There remain two parameters to specify, related to the microphysics of the gas. These are the Doppler parameter b and the ortho- to para-hydrogen ratio. Both of these must be regarded as rather uncertain. We consider two sets of extreme values. For the uniform sphere model, we assume $b = b_{\text{min}}$ and an ortho- to para- ratio of zero (ie. no ortho-hydrogen), while for the truncated isothermal sphere model, we assume $b = b_{\text{max}}$ and an ortho- to para- ratio of three.

[†] This does not apply to very dense gas, where H_2 forms predominantly via three-body processes and the H_2 fractional abundance is typically near unity.

Our two models now give strong upper and lower limits on t_{dis} .

Applying our calculations to the concrete example of a halo of gas mass $M = 10^5 M_{\odot}$, with an H₂ abundance of 5×10^{-4} that formed at $z = 30$, we find that

$$8.3 \times 10^4 \text{ yr} < t_{\text{dis}} < 9.5 \times 10^5 \text{ yr}. \quad (40)$$

To put these figures into context, the massive stars we are considering will have main sequence lifetimes that are typically of the order of a few Myr, so in this particular case cooling will cease before the end of the star's life.

We next proceed to vary the parameters individually, to examine their relative importance. We take as a fiducial model the example above, vary M , x_{H_2} and z over their allowed range, and plot the results in figures 4 to 6 respectively. Several points are immediately apparent.

Firstly, the most important parameter for determining t_{dis} is the mass of gas in the halo. Its importance is due to the weak dependence of N_{H_2} upon mass. As the mass of the halo increases, the mass of H₂ within it increases proportionately, but the column density increases at a much slower rate. This means that a large increase in mass leads to only a small increase in the amount of radiation absorbed, and hence the increase in t_{dis} is large. Moreover, our constraint upon the halo mass is much poorer than those upon redshift or H₂ abundance.

Another conclusion that can be drawn is that changes in z (which are equivalent to changing the H₂ column density while keeping the *amount* of H₂ fixed) have only a small effect. This is because the H₂ column densities for the halos considered generally correspond to equivalent widths in the square-root portion of the curve of growth. Thus, large changes in N_{H_2} lead to only small changes in f_{abs} and t_{dis} .

It is clear from figure 4 that in the smallest halos photodissociation equilibrium will be reached quickly, within a small fraction of the stellar lifetime. We can therefore apply the results of Omukai & Nishi with confidence to such halos. As the mass of gas in the halo increases, however, t_{dis} becomes steadily more important, and the assumption of equilibrium breaks down. For example, when the gas mass is $10^6 M_{\odot}$, the *lower* limit on t_{dis} is approximately 10^6 years – a significant fraction of the stellar lifetime. We cannot safely apply Omukai & Nishi's results to such halos, and they may undergo significant cooling right up until the point at which the massive star becomes a supernova.

This situation changes somewhat if we increase N_{dis} , as would occur if a more massive star formed, or if several massive stars formed at approximately the same time. In this case our estimated values of t_{dis} would be correspondingly lower, and cooling would be suppressed much sooner. Nevertheless, even in this case we cannot conclude that UV feedback necessarily implies a low star formation efficiency, as much gas may already have cooled and collapsed into dense clumps by the time the first star forms. To properly ascertain the effects of the ultraviolet radiation on star formation efficiency, we must also consider the fate of H₂ within these dense clumps.

4.2 Dense clumps

The second component of our protogalactic model consists of dense clumps of gas embedded in the diffuse halo. The

Figure 4. Upper and lower limits on the photodissociation timescale, plotted as a function of the mass of gas in the halo, for halos forming at $z = 30$ with H₂ abundance $x_{\text{H}_2} = 5 \times 10^{-4}$.

Figure 5. As figure 4, but examining the effects of varying the redshift of formation. The gas mass is fixed at $10^5 M_{\odot}$, and the H₂ abundance is again $x_{\text{H}_2} = 5 \times 10^{-4}$.

free-fall timescale of such clumps, t_{ff} , will be considerably shorter than that of the halo. In the absence of significant pressure support, we would expect the clumps to collapse quickly and presumably to form stars on a similar timescale. This will only be true so long as the clumps remain able to cool efficiently. We assume that clumps will collapse and form stars only if they retain their H₂ until collapse, ie. only if $t_{\text{dis}} > t_{\text{ff}}$.

To see whether clumps are capable of preserving their H₂ for long enough to form stars, we need to calculate both of these timescales. The free-fall timescale is straightforward; for gas of density ρ it is simply

$$\begin{aligned} t_{\text{ff}} &= (G\rho)^{-1/2} \\ &\simeq 10^8 n^{-1/2} \text{ yr}. \end{aligned} \quad (41)$$

Figure 6. As figure 4, but examining the dependence upon H₂ abundance. The gas mass is $10^5 M_{\odot}$ and the halos form at a redshift $z = 30$.

Comparing this with a typical massive star lifetime of a few Myr, we see that we need not consider clumps with densities lower than 10^3 cm^{-3} ; at such densities, clumps will not have collapsed by the time the star becomes a supernova. Their subsequent evolution will be shaped by their interaction with the supernova remnant, a topic which lies beyond the scope of this paper.

To calculate t_{dis} for denser clumps we first note that $f_{\text{inc}} \ll 1$; individual clumps receive only a small fraction of the total flux from the star. Thus, despite there being significantly less H_2 in a clump than in the halo, it does not follow that clumps will have a shorter dissociation timescale.

For a clump of angular size Ω , as seen from the star, we have $f_{\text{inc}} = \Omega/4\pi$. In the particular case of a spherical clump, in the small angle approximation,

$$f_{\text{inc}} = \frac{R^2}{4D^2}, \quad (42)$$

where R is the radius of the clump, and D is the distance from clump to star. For simplicity we do not consider ellipsoidal or filamentary clumps; however, it would be straightforward to extend our analysis to them.

Modelling the density structure of the clumps is problematic; we have little theoretical guidance. Consequently, we adopt the somewhat unrealistic model of a uniform density sphere, with Doppler parameter $b = b_{\text{max}}$ and an ortho-to para-hydrogen ratio of three. These choices will minimize t_{dis} and maximize t_{ff} for a given clump, inasmuch as this is compatible. We can then be confident that if our model clumps collapse, so would more realistic clumps.

For a given clump we specify four parameters: its mass (in units of M_{\odot}), density, H_2 abundance and distance from the star. From these quantities, we calculate the clump radius

$$R = 6.2 \times 10^{18} M^{1/3} n^{-1/3} \text{ cm}, \quad (43)$$

and the H_2 column density along a radial ray passing through the centre of the clump,

$$N_{\text{H}_2} = 1.2 \times 10^{19} x_{\text{H}_2} M^{1/3} n^{2/3} \text{ cm}^{-2}. \quad (44)$$

From these we calculate f_{inc} , f_{abs} and f_{dis} and thence t_{dis} . We find that

$$t_{\text{dis}} = 20 x_{\text{H}_2} M^{1/3} n^{2/3} D^2 \left(\frac{N_{\text{dis}}}{10^{48}} \right)^{-1} f_{\text{abs}}^{-1} f_{\text{dis}}^{-1} \text{ yr}, \quad (45)$$

where D is in parsecs.

We immediately see that the dependance of t_{dis} upon clump mass is weak, unlike the case of the halo. This is due to the scaling of f_{inc} with clump mass; since massive clumps are larger, they absorb more of the flux, which, to a great extent, offsets the fact that they contain more H_2 . This weak dependence is rather fortunate, as we know very little about the mass spectrum of clumps to be found within protogalaxies.

The dependence of t_{dis} upon x_{H_2} is stronger, but we do not expect the H_2 abundance in clumps to differ much from that in the halo; consequently, the same tight constraints apply, and the overall uncertainty in t_{dis} is small. In very dense clumps, as noted previously, this does not apply – three-body formation of H_2 allows much higher abundances to form. However, this process only becomes important above a number density of 10^8 cm^{-3} ; as we show below, in clumps of this density, H_2 will in any case survive until collapse.

Figure 7. The distance at which the dissociation and free-fall timescales are equal, plotted as a function of clump density, for a clump with $x_{\text{H}_2} M = 10^{-4}$.

Figure 8. As figure 7, but for a clump with $x_{\text{H}_2} M = 10^{-3}$.

Figure 9. As figure 7, but for a clump with $x_{\text{H}_2} M = 10^{-2}$.

The most important parameters for determining the fate of a clump are thus its density and its distance from the star. For a given mass and H_2 abundance, we can set t_{dis} equal to t_{ff} and solve for the critical distance D_{crit} at which H_2 survives until collapse; this will be a function of clump density. We examine three cases, with different values of M and x_{H_2} . Rather than varying these quantities separately, we instead vary the combination $x_{\text{H}_2} M^{1/3}$, as it is in this form that the mass and H_2 abundance appear in our expressions for the H_2 column density and for t_{dis} . We examine cases corresponding to values of 10^{-4} , 10^{-3} and 10^{-2} for $x_{\text{H}_2} M^{1/3}$; we plot the results in figures 7, 8 and 9 respectively. These values cover the range of clump masses and H_2 abundances that we would reasonably expect to encounter.

The qualitative behaviour is clear. High density clumps easily preserve their H_2 until collapse, unless they happen to be very close to the star. Lower density clumps, on the other hand, are prevented from collapsing at significant distances from the star – to put these values for D_{crit} into context, the virial radius of a $10^6 M_{\odot}$ halo collapsing at $z = 30$ is about 100 pc.

To quantify the effects of ultraviolet feedback upon star formation within clumps is rather harder, as in this case we are not interested in the absolute value of D_{crit} so much as its size relative to the size of the star-forming region. Nevertheless, a rough guide to the expected effects would come from assuming that stars will form only in clumps denser than some threshold, say $n = 10^6 \text{ cm}^{-3}$, for which D_{crit} is small compared to the size of the protogalaxy. If we increase N_{dis} (by forming more massive stars, for example), then this density threshold will increase. On the other hand, we expect our model to give us an upper limit on D_{crit} and hence on the density threshold; realistic clumps may well survive much nearer to massive stars than our model clumps do.

5 COMPLICATIONS

In order to construct a simple approximation, we have omitted a number of complicating features. Some of these omissions call for further comment. These are the heating of the gas during photodissociation, the effects of circumstellar material, the contribution to H₂ dissociation made by H II regions and stellar winds and the possible effects of dust.

5.1 Gas heating due to photodissociation

Rather than simply suppressing cooling, H₂ photodissociation will in actual fact lead to heating of the gas, as some of the energy of the Lyman-Werner photon is transferred to the kinetic energy of the hydrogen atoms. However, it is easy to show that the effects of this heating upon our protogalactic model will be small. A single photodissociation typically produces 0.4 eV of heat (Black & Dalgarno 1977), and thus destruction of all of the H₂ within a cloud will increase the temperature (in the absence of any other heating or cooling) by

$$\begin{aligned} \Delta T &= \frac{6.4 \times 10^{-12}}{k} x_{\text{H}_2} \text{ K} \\ &\simeq 50000 x_{\text{H}_2} \text{ K}. \end{aligned} \quad (46)$$

As our upper limit on x_{H_2} is 10^{-3} , this corresponds to a maximum temperature increase of roughly 50 K. This is too small to significantly affect our results, and thus we are justified in ignoring this effect.

5.2 Circumstellar material

Since we expect stars to form within dense clumps of gas, it is reasonable to assume that they will still be surrounded by large amounts of molecular gas after they have formed. If this is so, then this gas may absorb most of the Lyman-Werner photons emitted by the star, in which case dissociation in the halo or in other dense clumps will be reduced until the bulk of the circumstellar H₂ is destroyed.

If this material is distributed in a spherically symmetric fashion around the newly formed star, then it is quite easy to estimate its effects. There are two possibilities. If the bulk of the gas is of moderate density, then we can simply adapt our estimate of t_{dis} for a spherically symmetric halo to the case of a far less massive clump. In this case, it is clear that

the dissociation timescale will be short. On the other hand, if the circumstellar gas is dense enough to form H₂ via three body reactions, then $x_{\text{H}_2} \simeq 1$, and the amount of H₂ gas can be large. However, such a high density also implies $f_{\text{abs}} \simeq 1$, which goes some way to compensating for this. We find that:

$$t_{\text{dis}} \simeq 20M \left(\frac{N_{\text{dis}}}{10^{48}} \right)^{-1} \text{ yr}. \quad (47)$$

This is short compared to the stellar lifetime for any plausible value of the clump mass.

However, both of these cases are rather unrealistic. Not only do they neglect the effects of ionizing radiation and stellar winds, which can be expected to help clear away any surrounding molecular gas, but they also assume spherical symmetry. In reality, we would expect the gas surrounding newly formed primordial stars to be highly flattened. This is necessary for efficient outward transportation of angular momentum, in the assumed absence of magnetic fields. In this case, we would expect the circumstellar gas to survive for a much greater time, as it will receive less of the stellar flux. Nevertheless, it will not seriously affect our estimates as it will shield only the small fraction of gas which lies in its shadow.

5.3 H II regions and stellar winds

Both ionizing radiation and stellar winds offer plausible alternative mechanisms by which H₂ may be destroyed, and which may reduce the importance of H₂ photodissociation. It is thus important to examine their effects.

Stellar winds from low metallicity stars have recently been studied by Kudritzki (2000). They are found to be several orders of magnitude less efficient than winds from solar metallicity stars, due to the lack of metal lines with which to drive them. It thus seems reasonable to neglect their effects.

Ionizing radiation, on the other hand, is rather more important. Ionizing flux from a massive star will lead to the formation of an H II region surrounding the star, whose size and subsequent evolution depend upon the density distribution near the star. A general treatment of the formation and evolution of this H II region requires solution of the coupled radiative transfer and hydrodynamics problem, which is clearly impractical. However, we can get some guidance as to the expected effects by examining some relevant cases where the high degree of symmetry allows for analytical solution.

The simplest case is that in which the gas surrounding the massive star is spherically symmetric and of uniform density. In this case, the growth of the H II region can be divided into two phases (Yorke 1986). During the first phase, an R-type ionization front sweeps into the surrounding gas at supersonic velocity. As the radius of this front approaches the Strömgren radius

$$r_S = \left(\frac{3N_{\text{ion}}}{4\pi n^2 \alpha_B} \right)^{1/3}, \quad (48)$$

(where N_{ion} is the number of ionizing photons produced by the source per second and α_B is the case B recombination coefficient), it slows and changes to a D-type front. As it does so, a shock front forms and detaches from the ionization front, preceding it into the surrounding gas. The sub-

sequent expansion phase leads to the formation of a dense shell of H I gas between the shock and the ionization front. The thickness of this shell is small compared to the radius of the H II region. The latter is given approximately by (Yorke 1986)

$$r_I = r_S \left[1 + \frac{7}{4} \frac{c_s(t - t_0)}{r_S} \right]^{4/7}. \quad (49)$$

Here c_s is the isothermal sound speed in the H II region, and $t - t_0$ is the time since the formation of the shock. The expansion phase ends when the H II region attains pressure equilibrium with the surrounding gas. This occurs at a radius

$$r_f \simeq \left(\frac{2T_2}{T_1} \right)^{2/3} r_S, \quad (50)$$

where T_1 and T_2 are the temperatures of the H I and H II gas respectively. For gas cooled by H_2 , the former will be approximately 200 K (or smaller if HD cooling is important), while the latter will be approximately 10^4 K; thus $r_f \simeq 20r_S$. However, unless the gas density is extremely large, the main sequence lifetime of the central star will generally be too short to allow the H II region to reach this pressure equilibrium; in this case, the maximum size of the H II region is given by r_I , evaluated at the end of the star's life.

If, rather than a uniform density distribution, we have an instead an inhomogeneous distribution, with uniform *mean* density $\langle n \rangle$, then the above calculations still give us a reasonable estimate for the size of the resulting H II region, as long as we replace n^2 by $\langle n^2 \rangle$. We generally express this by means of a clumping factor C such that $C = \langle n^2 \rangle / \langle n \rangle^2$. This results in a value for r_S a factor $C^{1/3}$ smaller than in the uniform density case.

The other simple case we examine is the propagation of an H II region into a radially decreasing density distribution. This problem has been studied by Franco, Tenorio-Tagle & Bodenheimer (1990), who show that if the density drops off more sharply than $n \propto r^{-3/2}$ then the resulting H II region is unbounded by recombination. This implies that the entire halo may become ionized; however, as this requires at least one ionizing photon per hydrogen atom (assuming that secondary ionization by energetic photons is negligible), this cannot take place on a timescale less than

$$\begin{aligned} t_{\text{ion}} &= \frac{M}{\mu m_H N_{\text{ion}}} \\ &= \frac{3.2 \times 10^{49} M}{N_{\text{ion}}} \text{ yr.} \end{aligned} \quad (51)$$

where M is the mass of gas in the halo.

To apply these calculations to our model halos, we need to specify N_{ion} . We can calculate this in a similar fashion to N_{dis} , by modelling the star as a black body. Whilst this will inevitably be inaccurate, the values we obtain compare well with those from the more detailed treatment of Tumlinson & Shull (1999). For the case of a $25 M_\odot$ star considered previously, we find that $N_{\text{ion}} \simeq 7 \times 10^{48}$. If we take the temperature of the H II region to be 10^4 K, we then have

$$r_S = 60n^{-2/3} \text{ pc}, \quad (52)$$

where we have used the case B recombination coefficient from Osterbrock (1989).

For our uniform density halo model, this implies

Figure 10. The final size of the H II region formed by a $25 M_\odot$ star in halos of total mass $10^5 M_\odot$ (dotted line) and $10^8 M_\odot$ (dashed line) that is predicted if the density is assumed constant and equal to the central density of a truncated isothermal sphere of the same mass, forming at the specified redshift. It is plotted in units of r_0 , the core radius of the truncated isothermal sphere. Also plotted is the radius at which the density profile of the sphere begins to fall off more steeply than $r^{-3/2}$ (solid line).

$$r_S = \frac{8.1 \times 10^4}{(1+z)^2} \text{ pc}. \quad (53)$$

However, given the unrealistic nature of this density profile, this figure is probably not relevant. For the more realistic truncated isothermal sphere model, we have

$$r_S = \frac{3.8 \times 10^3}{(1+z)^2} \text{ pc}, \quad (54)$$

assuming that the H II region is confined to the core of the density profile. From this we can calculate r_I , given some value for $t - t_0$. We choose to adopt a value of 2 Myr, corresponding to a typical OB star lifetime.[†] In figure 10 we plot r_I (in units of the core radius r_0) for halos of total mass $10^5 M_\odot$ and $10^8 M_\odot$, at a range of redshifts. Also plotted is the radius at which the truncated isothermal sphere density profile becomes steeper than $r^{-3/2}$; this occurs at approximately $2.9r_0$.

For halos in which $r_I > r_0$, our analysis, which assumes constant density, begins to break down; for $r_I > 2.9r_0$, it breaks down completely, and the H II region will in reality become unbounded. We see from figure 10 that this is true for all but the most massive halos. Taking this at face value, it suggests that, as well as photodissociating H_2 throughout a halo, a single massive star will also ionize the entire halo. However, this prediction may owe much to the oversimplistic nature of our halo model. In a real halo, several factors are likely to combine to reduce our estimated values of r_I . Significant clumping is likely, and will increase the effective density and hence decrease r_I . Cooling will also increase the core density above that predicted by the truncated isothermal sphere model. Finally, non-thermal pressures due to infalling gas or turbulence may also restrict the expansion of the H II region.

Given these uncertainties, plus the rather ad hoc nature of the adopted density profile, the final size of the first H II regions is rather unclear. On the one hand, they may

[†] This is an overestimate, but will be reasonably accurate so long as the initial R-type phase is much smaller than the subsequent D-type phase.

be comparable in size to the halo; on the other, they may prove to more closely resemble the ultracompact H II regions seen in some local molecular clouds. To decide between these scenarios will require a better understanding of the density and velocity structure of the first protogalaxies than we currently possess, a deficiency that high resolution numerical simulations are beginning to address (Abel et al. 2000).

Finally, it should be noted that even if the final H II region is large, this will significantly affect our conclusions only in the smallest halos. In larger halos, $t_{\text{dis}} < t_{\text{ion}}$, so cooling is still regulated in the first instance by photodissociation. Similarly, clumps dense enough to preserve their H₂ until collapse will in general have $t_{\text{ff}} < t_{\text{ion}}$, and thus will only be affected by the H II region if they lie close to the massive star.

5.4 Dust

Dust, if present, can affect our simple picture in three ways. Firstly, H₂ can form by direct radiative association on grain surfaces. The rate of H₂ formation by this process is independent of ionization. Consequently, at low temperatures the argument of Nishi & Susa (1999) is no longer valid, and the H₂ abundance may be significantly larger than 10^{-3} .

Second, dust absorbs strongly in the Lyman-Werner bands, and will reduce the number of photons able to cause photodissociation.

Finally, dust implies the presence of metals, which may replace H₂ as the dominant coolant at low temperatures. This means that H₂ dissociation need not imply cessation of star formation, if the abundance of metals and dust is high enough to cool the gas.

With each possible effect, we associate a critical metal abundance, above which the effect becomes important. We estimate these below.

5.4.1 H₂ formation on grains

H₂ forms on grain surfaces in the local ISM at a rate (Draine & Bertoldi 1996)

$$R_{\text{dust}} = 6.0 \times 10^{-18} T^{1/2} n_{\text{H}}^2 \text{ cm}^{-3} \text{ s}^{-1}, \quad (55)$$

where n_{H} is the H I number density. If we assume both that dust in low metallicity gas has the same distribution of grain sizes as that in the Milky Way, and that the ratio of dust to metals is independent of metallicity, then the rate of H₂ formation on grains in metal-poor gas is simply

$$R_{\text{dust}} = 6.0 \times 10^{-18} T^{1/2} \left(\frac{Z}{Z_{\odot}} \right) n_{\text{H}}^2 \text{ cm}^{-3} \text{ s}^{-1}. \quad (56)$$

Both of these assumptions are somewhat *ad hoc* but they are the simplest that we can make, and should give us a reasonable idea of the importance of dust.

Comparing R_{dust} with the rate of the H⁺ gas phase reaction (equation 8), we see that at low metallicities the gas phase reaction will dominate, while at higher metallicities formation on grain surfaces dominates. A reasonable definition of the critical metallicity is that at which the two rates are equal. This is given by

$$Z_{\text{crit}} = \frac{T^{1/2} x_{\text{e}}}{6} Z_{\odot}. \quad (57)$$

For a protogalaxy with temperature $T = 10^3$ K and ionization $x_{\text{e}} = 3 \times 10^{-4}$ (corresponding to the residual abundance from recombination), this implies that

$$Z_{\text{crit}} = 1.6 \times 10^{-3} Z_{\odot}. \quad (58)$$

Thus, H₂ formation on dust may become important at metallicities as small as a few thousandths of the solar value.

5.4.2 Extinction

To assess the importance of dust extinction, we calculate the optical depth due to dust at 1000 Å. If we again assume that primordial dust is similar to that in the Milky Way, and that the dust to metal ratio is constant, then the optical depth is given by

$$\tau_{\text{d}} = \sigma_{\text{d},1000} N_{\text{H}} \left(\frac{Z}{Z_{\odot}} \right), \quad (59)$$

where $\sigma_{\text{d},1000}$ is the absorption cross-section at 1000 Å, N_{H} the column density of hydrogen, again assuming a constant dust to metal ratio. Setting $\tau_{\text{d}} = 1$, and adopting a value of $2 \times 10^{-21} \text{ cm}^2$ for $\sigma_{\text{d},1000}$ (Draine & Bertoldi 1996), we can solve for the critical metallicity at which the dust becomes optically thick. We find that

$$Z_{\text{crit}} = 5 \times 10^{20} N_{\text{H}}^{-1} Z_{\odot}. \quad (60)$$

The critical metallicity thus depends upon the column density, and consequently upon the properties of the individual halo or clump. The range of column densities included in our model is large, being approximately 10^{19} to 10^{23} cm^{-2} for halos and 10^{21} to 10^{25} cm^{-2} for clumps. Adopting the median values of 10^{21} cm^{-2} and 10^{23} cm^{-2} as representative of ‘typical’ halos and clumps, this implies critical metallicities of $0.5 Z_{\odot}$ and $5 \times 10^{-3} Z_{\odot}$ respectively. This suggests that dust absorption in halos is unlikely to be important until substantial enrichment has occurred, while clumps become shielded at much smaller metallicities.

5.4.3 Metal line cooling

The presence of metals makes available a large number of alternative cooling processes, such as fine structure emission from C II or O I. At a sufficiently high metallicity, these dominate the low temperature cooling, rendering the photodissociation of H₂ irrelevant. To properly treat their effects calls for a detailed model of the chemistry and ionization balance in the gas. However, we can gain a simple estimate of the metallicity at which metal line cooling becomes important by considering cooling from a single representative process. In highly opaque gas, where the UV flux is small, Tielens & Hollenbach (1985) find that the dominant coolant is CO. However, this is even easier to dissociate than H₂ and thus can play no role in the cooling of the gas after H₂ photodissociation. In gas where the UV flux is high, on the other hand, the dominant coolant is the O I 63 μm fine structure line. This cools the gas at a rate

$$\Lambda = 9.8 \times 10^{-29} T^{0.67} \exp \left(\frac{-228}{T} \right) \frac{Z}{Z_{\odot}} n_{\text{H}}^2 \text{ erg cm}^{-3} \text{ s}^{-1}, \quad (61)$$

where we have adopted a value of 7.4×10^{-4} for the O I abundance in solar metallicity gas (Pagel 1997). Equating this to

Figure 11. The metallicity at which the cooling rate due to the O I 63 μ m fine structure line is equal that due to molecular hydrogen. We assume an oxygen abundance of 7.4×10^{-4} in solar metallicity gas and an H₂ abundance of 10^{-3} .

the low density H₂ cooling rate from Galli & Palla (1998), we solve for the critical metal abundance as a function of temperature, assuming an H₂ abundance of 10^{-3} . We plot our results in figure 11. We see that at very low temperature, a small O I abundance makes a significant difference. On the other hand, as the temperature increases, the effects of O I cooling steadily decrease. For gas at a temperature of 1000 K (a typical virial temperature), we find that $Z_{\text{crit}} = 0.07 Z_{\odot}$. This is clearly an upper limit; including other metal lines, or reducing the H₂ abundance will decrease it. A value of $Z_{\text{crit}} \simeq 0.01 Z_{\odot}$ would therefore seem to be a reasonable order of magnitude estimate.

6 CONCLUSIONS

In this paper we have examined the destruction of H₂ within primordial protogalaxies by ultraviolet radiation from the first massive stars to form within them. To do this we have constructed a simple analytical model of photodissociation which incorporates the essential physics and which should give a reasonable order of magnitude estimate of the time taken to destroy the bulk of the H₂. Using this model, we have shown that for diffuse gas in small protogalaxies the conclusions arrived at by Omukai & Nishi are broadly correct: a single massive star will produce sufficient ultraviolet radiation to suppress H₂ cooling within its own lifetime. On the other hand, in larger protogalaxies the photodissociation timescale becomes comparable to the stellar lifetime, and the effects on gas cooling are correspondingly smaller.

We have also examined the effects of photodissociation upon dense clumps of gas, such as we expect to find in star forming regions. Using a simplified clump model, we find that there is a density-dependent critical distance, D_{crit} , beyond which the photodissociation timescale is longer than the clump free-fall timescale. For moderate density clumps, D_{crit} is comparable to the dimensions of the star-forming region, but it decreases sharply as clump density increases. Since clumps can continue to collapse only if cooled by H₂, it is reasonable to assume that the survival of H₂ is a prerequisite for star formation. Thus, stars will form only in clumps that lie sufficiently far from any massive star. An alternative way of considering this is to suppose that stars will only form in clumps above some density threshold, where this is chosen such that the corresponding D_{crit} is much smaller

than the dimensions of the protogalaxy. In this case, an appropriate value for the density threshold is approximately 10^6 cm^{-3} , although this is at best a rough estimate.

We have also examined a number of factors which may complicate this simple picture. The most important prove to be ionizing radiation and the presence of dust and metals.

The effects of ionizing radiation are sensitive to the small-scale density distribution within the halo, and are thus hard to assess accurately. However, we have shown that even if there is sufficient flux to ionize most or all of the halo, this will generally take longer than H₂ photodissociation. Consequently, it should have little effect on our central results.

Dust and metals can lead to more significant changes in the physics, but these begin to take effect only once the metallicity of the gas exceeds some critical value, which we estimate as $Z_{\text{crit}} \sim 10^{-3} Z_{\odot}$.

Our overall picture of star formation in the first protogalaxies is thus one in which the formation of dense clumps plays a central role. Ultraviolet feedback is effective at suppressing cooling in diffuse gas (to a greater or lesser degree), but has little effect on dense clumps. Moreover, while it is likely that supernovae will eventually remove all of the gas from the system, we expect this to occur *after* these dense clumps have collapsed. An upper limit to the star formation efficiency is thus set by the fraction of gas that has formed into dense clumps by the time that the first massive stars ignite. This may well be small, but it is too early to conclude this with any real certainty, and thus scope for significant star formation within the first cosmological objects still exists.

REFERENCES

- Abel T., Bryan G. L., Norman M. L., 2000, preprint, astro-ph/0002135
- Abel T., Anninos P., Zhang Y., Norman M. L., 1997, *New A.*, 2, 181
- Abgrall H., Roueff E., 1989, *A&AS*, 79, 313
- Abgrall H., Bourlot J. L., des Forêts G. P., Roueff E., Flower D. R., Heck L., 1992, *A&A*, 253, 525
- Allison A. C., Dalgarno A., 1969, *At. Data*, 1, 91
- Black J. H., Dalgarno A., 1977, *ApJS*, 34, 405
- Blitz L., Williams J. P., 1997, *ApJ*, 488, L145
- Bromm V., Coppi P. S., Larson R. B., 1999, *ApJ*, 527, L5
- Ciardi B., Ferrara A., Governato F., Jenkins A., 1999, preprint, astro-ph/9907189
- Cojazzi P., Bressan A., Lucchin F., Pantano O., 1999, preprint, astro-ph/9908030
- de Jong T., 1972, *A&A*, 20, 263
- Draine B. T., Bertoldi F., 1996, *ApJ*, 468, 269
- Elmegreen B. G., 1997, *ApJ*, 477, 196
- Ferrara A., Tolstoy E., 1999, preprint, astro-ph/9905280
- Franco J., Tenorio-Tagle G., Bodenheimer P., 1990, *ApJ*, 349, 126
- Galli D., Palla F., 1998, *A&A*, 335, 403
- Haiman Z., Abel T., Rees M. J., 1999, preprint, astro-ph/9903336
- Haiman Z., Thoul A. A., Loeb A., 1996, *ApJ*, 464, 523
- Kudritzki R. P., 2000, in Weiss A., Abel T., Hill V., *The First Stars: Proceedings of the second MPA/ESO workshop*, Springer, Heidelberg
- MacLow M.-M., Ferrara A., 1999, *ApJ*, 513, 142
- Nakamura F., Umemura M., 1999, *ApJ*, 515, 239
- Nishi R., Susa H., 1999, *ApJ* 523, L103

- Omukai K., Nishi R., 1999, *ApJ*, 518, 64
- O’Neil S. V., Reinhardt W., 1978, *J. Chem. Phys.*, 69, 2126
- Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, University Science Books, California
- Pagel B. E. J., 1997, *Nucleosynthesis and Chemical Evolution of Galaxies*, Cambridge Univ. Press, Cambridge
- Rees M. J., Ostriker J. P., 1977, *MNRAS*, 179, 541
- Rodgers C. D., Williams A. P., 1974, *J. Quant. Spectrosc. Rad. Transf.*, 14, 319
- Shapiro P. R., Iliev I. T., Raga A. C., 1999, *MNRAS*, 307, 203
- Stecher T. P., Williams D. A., 1967, *ApJ* 149, L29
- Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, *ApJ*, 474, 1
- Tielens A. G. G. M., Hollenbach D., 1985, *ApJ*, 291, 722
- Tumlinson J., Shull J. M., 1999, preprint, astro-ph/9911339
- Yorke H. W., 1986, *ARA&A*, 24, 49